

Achieving Cloud Security using Third Party Auditor and Preserving Privacy for Shared Data Over Public Cloud

Chakresh Kumar¹ and Dr. Annamalai Giri A. A²

Research Scholar, Faculty of Computer Science, Department of Science, Arunodaya University, India¹

Research Guide, Faculty of Computer Science, Department of Science, Arunodaya University, India²

Abstract: *Cloud computing is an emerging technology that will receive more attention in the future from industry and academia. The cost of this technology is more attractive when it is compared to building the infrastructure. However, there are many security issues coming with this technology as happens when every technology matures. In this research paper data security, data integrity and access control in the public cloud is achieved with significant results. In this process, Third Party Auditor (TPA) and user separation are used successfully. The TPA has a hybrid algorithm for signature generation called MD5withRSA. The access control is used for separate users from data owners and only those users can have access to the owner's data who have granted access by data owners. Data is compressed without affecting the quality of data to reduce the storage cost. The compressed data is then stored in chunks to provide security. The proposed system can be further extended to improve the TPA performance of different types of data on cloud environment.*

Keywords: Cloud computing, TPA, Public Cloud, Data Owner

I. INTRODUCTION

Cloud computing enabled distributed data storage and at the same time reduced the usage costs. Cloud facilitated access to data anytime from anywhere and also from multiple locations. The user is relieved of the complexities of hardware and software needs for data storage and sharing mechanisms and enjoys data -storage location independence. The user can simply utilize the services provided by the different organizations offering cloud services for data storage and access. With ease of access and storage on cloud, the issues of usage authorization and data security pop up. Typically, authorized users, are allowed to access cloud storage with restrictions at different levels, with the access control being managed by the system administrator. One of the restrictions that help securing data is providing access to usage of data for a particular use and to restrict the user's number approaches that can be used for access and usage. Attributes are the deciding factor for some of the user level grouping formation. Content storage and content sharing for different purposes being the major use of cloud, data security while transferring or storing the file is an area of concern requiring efficient solutions. Amazon simple storage services and amazon elastic compute cloud have enabled users from different segments like entertainment, legal, product development and business organizations to store and access different types of data. The process of maintaining data security is generally not a concern for the user once the system is driven by the cloud with the cloud security model for data being in place. Yet, security concerns remain a higher priority for the data on the cloud in terms of external attacks on the cloud data that corrupts the content and violates the integrity of cloud storage. As the correctness of user data is not maintained by the cloud security systems, the users would need mechanisms to address the data security in the cloud system These mechanisms should not require the users of cloud to manage the data at frequent intervals. In this paper, we propose a method that includes cryptography mechanism with significant modifications. In addition, a third party auditor (TPA) is introduced to monitor/ keep watch on the system. The proposed method offers data confidentiality and integrity for the data in the cloud.

TPA checks integrity and secure storage of data in the cloud. The auditing proof generation can be generated in two different mechanisms: first, the user carries out the procedure to verify the correctness of data that is stored in the cloud. That is, the cloud data owner checks the data - leading to a time and cost burden for user every time a check is carried

out. Alternately, in the second method, the TPA verifies the data for secure storage for all the users in a single instance. This batch processing method helps TPA execute the procedure for all users in a single turn. The TPA proposed can check data for many users in a single instance and generate a report for all users. The method proposed in this work for cloud data security uses multiple data security attributes as described in section III. The level of security is further checked by the outside agent TPA to keep watch on every step. This process ensures that integrity of data is maintained and, also, every user's data correctness is maintained. In addition, authentication is provided in the system to allow only authorized users, the access to stored data on the cloud. The proposed method offers efficiency with the addition of TPA supporting dynamic and batch processing of auditing of data for many users.

1.1 Problem Definition

The cloud providers may sometimes face technical outages that can happen due to various reasons, such as loss of power, low Internet connectivity, data centers going out of service for maintenance, etc. The Existing work there is no any scope for constructive and flexible data-sharing framework with a third-party auditor approach. There is no any integrity of the user and user's data on cloud while different operations performing in the environment.

1.2 Background

Boyang Wang & Baochun Li & Hui Li (2014) proposed for privacy-preserving for public data. The data is shared across all users of that cloud the method known as the Oruta method. The public verifier has no right to detect who is the signer on each block yet method can perform the auditing function without all attributes. In the system, the author has used a ring signature for the creation of a homomorphic authenticator on the cloud. The system can perform the batch auditing process on data for fast processing. The authors mentioned two problems that they will study in the future; the first one is traceability that is the original user can reveal the identity in some special situations. The method also supports the data freshness on the cloud while processing a different request from different users. It helps to preserve data identity throw-out the complete process which helps to achieve privacy.

Cong Wang, Qian Wang, Kui Ren, Wenjing Lou (2013) tried to provide a TPA to tackle this problem. In proposed method attributes of data and the dependency stored at the cloud on different locations is discussed in detail. The author has used a homomorphic linear authenticator for the cloud data storage and random masking technique for TPA. This method helps that TPA doesn't have any knowledge about the cloud-stored data. Even though without knowledge of data TPA perform auditing task efficiently. This separation helps in avoiding the data leakage problem on cloud storage. The paper states that batch auditing can be introduced to increase the process execution with efficiently.

KaipingXue, Weikeng Chen, Wei Li, Jianan Hong, Peilin Hong (2018) has proposed two different methods for cloud data storage. One method with the data being stored at the cloud. The second is with data owner where they have used access control mechanism, with the encryption method. The proposed framework not only provides resource consumption but also protects from the DDoS/EDoS attacks in the cloud environment. The method used in this paper is CP-ABE which helps to keep data secure from the malicious attacks during the cloud operations and presence over it. The author has used bloom filter and probabilistic check to reduce resource overhead. The results at the end show that the task is achieved successfully and overheads are very less with comparative other methods.

Cong Wang, & Bingsheng Zhang & Kui Ren & Janet M. Roveda (2013) explained the method of image recovery in their work. The method is prominent with privacy concern. In the proposed method of OIRS where the compressed sensing method is used. The compression is a part of sensing data. The owners have the benefit of compressed sensing and also compression through the linear measurement process. The users in the cloud system use the cloud resources to outsourced the image and recover that particular image. The user gets an image without revealing the received compressed procedure or compressed samples or the content for recovering that image. OIRS is simple and efficient and can reconstruct images effectively. Users always pay more attention to data security in the cloud. In recent years, data integrity schemes have become one of the research hotspots. With the help of data integrity schemes, any data corruption or deletion can be discovered in time and then necessary measures can be taken to recover the data. To develop a better understanding of data integrity schemes, they carry out the relevant work from the audit model, soundness, and other aspects.

Wang et al. [2–7] noticed the problem of shared data integrity verification and proposed a public auditing method that supports efficient user revocation for shared data. To sum up, this scheme introduces proxy resignation technology to solve the problem. However, when the user is revoked, the cloud server is allowed to replace the previously signed data block of the revoked user to a legal group instead of the group member, which can cause efficiency problem. In addition, in scheme [8], the authors propose to enable efficient user revocation in identity-based cloud storage auditing for shared big data. On the other hand, Yu et al. [8, 9] proposed the issue of key security among cloud users. In these schemes, the key exposure in one-time period does not affect the security of cloud storage auditing in other time periods and verifiable out-sourcing of key updates.

Yu et al. [10] proposed that the storage and sharing services of cloud servers allow users to share data in the form of a group. As a group member, they have the right to view and modify shared data. Although users can easily share data, data integrity issues remain [11, 12]. Using TPA for public auditing results in the leakage of user's identity privacy [13]. Wang et al. [14] fully considered the confidentiality of the data in the public audit process and proposed a privacy scheme that used ring signature to protect group member. Adopting the ring signature can ensure that the TPA protects the user's identity privacy while verifying the integrity of the data. However, the efficiency of the scheme is reduced by the increasing number of team members. Meanwhile, the client also takes a lot of computing. Therefore, the scheme does not apply to large user groups.

Shen et al. [15] proposed a lightweight auditing scheme for shared data privacy protection, taking full account of the computational limitations of the resource constrained client. Using data blindness methods, the scheme allows (TPM) Third Party Medium instead of group users to sign the data. It not only reduces the burden on the client, but also ensures the privacy of identity during public auditing. Thus the identity of the data owner can be protected. However, this scheme does not support group dynamics and the traceability of data blocks.

Wang et al. [16] proposed another public audit method for sharing data privacy protection. Using dynamic broadcast technology, group members can be signed as the owner of the data when modifying the shared data, thereby protecting the privacy of the group members. It not only realizes the dynamic operation of data by group members, but also supports group dynamics. However, this scheme does not protect the identity of data owner, making the TPA steal the identity of the data owner during public auditing, and it does not support the traceability of data blocks.

The first method [17] allows only the data owner to audit. The second [18] method allows a third party auditor to audit. The audit process in both approaches is performed without retrieving the remote data. If only the data owner can verify the integrity of the outsourced data, then this scheme is considered to provide private audit ability. However, in some cases, it is not practically feasible for the data owner to remain online all time for data integrity verification. Hence, the data owner can delegate this responsibility for integrity verification to a third party auditor or other users. A data integrity scheme must have public audit ability property to support this audit delegation.

Data can be either static (backup or archival data) or dynamic nature (supporting operations like insertion, deletion, and modification). Providing integrity for dynamic data is more challenging than static data or just attaching data. Most of the schemes proposed in the literature are not able to handle dynamic data, such as the description of the schemes [19, 20] dynamic data handling characteristic demands that data integrity should remain intact, even after insertion, deletion, or modification.

In the schemes of Wang [21] and Zhang et al. [22], the soundness property of data integrity schemes ensures data reliability. Data integrity schemes are designed to prevent tampering. Therefore, if metadata is tampered with or corrupted intentionally or unintentionally by the CSP, this should be timely identified by a data integrity scheme. If the CSP can pass a challenge request without holding the data or with corrupted data, then a client will never be able to identify data corruption promptly, and the value of the data will be lost. Therefore, a good data integrity scheme requires that the server's response must be reliable.

Privacy protection should be emphasized in the process of data integrity verification. As involved in the scheme [23], privacy concerns are introduced due to public verifiability. On the premise that the data owner will not allow the disclosure of his private data to a third party auditor, the privacy preservation property demands that a third party auditor should not obtain any confidential information about the user's data but can still verify the integrity of outsourced data.

In the scheme [24], fairness means that a data integrity scheme should provide protection for an honest CSP against legitimate but dishonest users, who may attempt to accuse CSP of manipulating the outsourced data. If a data integrity scheme does not support fairness, it means dishonest users can damage CSP reputation.

II. CLOUD SECURITY FOR PRIVACY PRESERVING

Cloud security involves the procedures and technology that secure cloud computing environments against both external and insider cyber security threats. Cloud computing, which is the delivery of information technology services over the internet, has become a must for businesses and governments seeking to accelerate innovation and collaboration. Cloud security and security management best practices designed to prevent unauthorized access are required to keep data and applications in the cloud secure from current and emerging cyber security threats.

2.1 Cloud computing categories using Python

Cloud security differs based on the category of cloud computing being used. There are four main categories of cloud computing:

- **Public cloud services, operated by a public cloud provider:** These include software-as-a-service (SaaS), infrastructure-as-a-service (IaaS), and platform-as-a-service (PaaS).
- **Private cloud services, operated by a public cloud provider:** These services provide a computing environment dedicated to one customer, operated by a third party.
- **Private cloud services, operated by internal staff:** These services are an evolution of the traditional data center, where internal staff operates a virtual environment they control.
- **Hybrid cloud services:** Private and public cloud computing configurations can be combined, hosting workloads and data based on optimizing factors such as cost, security, operations and access. Operation will involve internal staff, and optionally the public cloud provider.

Here's a diagram showing common control plane across cloud models:

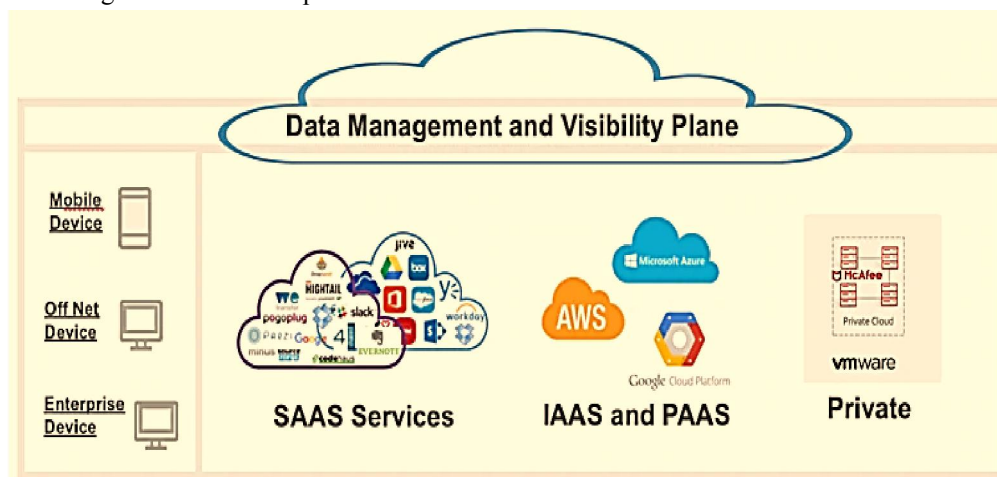


Figure 1. Common Control Plane Across Cloud Models

Source- <https://www.mcafee.com/enterprise/en-in/security-awareness/cloud.html>

When using a cloud computing service provided by a public cloud provider, data and applications are hosted with a third party, which marks a fundamental difference between cloud computing and traditional IT, where most data was held within a self-controlled network. Understanding your security responsibility is the first step to building a cloud security strategy.

2.2 Segmentation of Cloud Security Responsibilities

Most cloud providers attempt to create a secure cloud for customers. Their business model hinges on preventing breaches and maintaining public and customer trust. Cloud providers can attempt to avoid cloud security issues with the service they provide, but can't control how customers use the service, what data they add to it, and who has access.

Customers can weaken cyber security in cloud with their configuration, sensitive data, and access policies. In each public cloud service type, the cloud provider and cloud customer share different levels of responsibility for security. By service type, these are:

- **Software-as-a-service (SaaS)** — Customers are responsible for securing their data and user access.
- **Platform-as-a-service (PaaS)** — Customers are responsible for securing their data, user access, and applications.
- **Infrastructure-as-a-service (IaaS)** — Customers are responsible for securing their data, user access, applications, operating systems, and virtual network traffic. Within all types of public cloud services, customers are responsible for securing their data and controlling who can access that data. Data security in cloud computing is fundamental to successfully adopting and gaining the benefits of the cloud. Organizations considering popular SaaS offerings like Microsoft Office 365 or Sales force need to plan for how they will fulfil their shared responsibility to protect data in the cloud. Those considering IaaS offerings like Amazon Web Services (AWS) or Microsoft Azure need a more comprehensive plan that starts with data, but also covers cloud app security, operating systems, and virtual network traffic each of which can also introduce potential for data security issues.

2.3 Cloud Security Challenges

Since data in the public cloud is being stored by a third party and accessed over the internet, several challenges arise in the ability to maintain a secure cloud. These are:

- **Visibility into Cloud Data:** In many cases, cloud services are accessed outside of the corporate network and from devices not managed by IT. This means that the IT team needs the ability to see into the cloud service itself to have full visibility over data, as opposed to traditional means of monitoring network traffic.
- **Control Over Cloud Data:** In a third-party cloud service provider's environment, IT teams have less access to data than when they controlled servers and applications on their own premises. Cloud customers are given limited control by default, and access to underlying physical infrastructure is unavailable.
- **Access to Cloud Data and Applications:** Users may access cloud applications and data over the internet, making access controls based on the traditional data center network perimeter no longer effective. User access can be from any location or device, including bring-your-own-device (BYOD) technology. In addition, privileged access by cloud provider personnel could bypass your own security controls.
- **Compliance:** Use of cloud computing services adds another dimension to regulatory and internal compliance. Your cloud environment may need to adhere to regulatory requirements such as HIPAA, PCI and Sarbanes-Oxley, as well as requirements from internal teams, partners and customers. Cloud provider infrastructure, as well as interfaces between in-house systems and the cloud are also included in compliance and risk management processes.
- **Cloud-native Breaches:** Data breaches in the cloud are unlike on-premises breaches, in that data theft often occurs using native functions of the cloud. A Cloud-native breach is a series of actions by an adversarial actor in which they "land" their attack by exploiting errors or vulnerabilities in a cloud deployment without using malware, "expand" their access through weakly configured or protected interfaces to locate valuable data, and "ex-filtrate" that data to their own storage location.
- **Misconfiguration:** Cloud-native breaches often fall to a cloud customer's responsibility for security, which includes the configuration of the cloud service. Research shows that just 26% of companies can currently audit their IaaS environments for configuration errors. Misconfiguration of IaaS often acts as the front door to a Cloud-native breach, allowing the attacker to successfully land and then move on to expand and ex-filtrate data. Research also shows 99% of misconfigurations go unnoticed in IaaS by cloud customers. Here's an excerpt from this study showing this level of misconfiguration disconnect.

2.4 Cloud Security Solutions

Organizations seeking cloud security solutions should consider the following criteria to solve the primary cloud security challenges of visibility and control over cloud data.

Visibility into Cloud Data: A complete view of cloud data requires direct access to the cloud service. Cloud security solutions accomplish this through an application programming interface (API) connection to the cloud service. With an API connection it is possible to view:

- What data is stored in the cloud.
- Who is using cloud data?
- The roles of users with access to cloud data.
- Who cloud users are sharing data with.
- Where cloud data is located.
- Where cloud data is being accessed and downloaded from, including from which device.

Control Over Cloud Data: Once you have visibility into cloud data, apply the controls that best suit your organization. These controls include:

Data Classification: Classify data on multiple levels, such as sensitive, regulated, or public, as it is created in the cloud. Once classified, data can be stopped from entering or leaving the cloud service.

Data Loss Prevention (DLP): Implement a cloud DLP solution to protect data from unauthorized access and automatically disable access and transport of data when suspicious activity is detected.

Collaboration Controls: Manage controls within the cloud service, such as downgrading file and folder permissions for specified users to editor or viewer, removing permissions, and revoking shared links.

Encryption: Cloud data encryption can be used to prevent unauthorized access to data, even if that data is ex-filtrated or stolen.

Access to Cloud Data and Applications: As with in-house security, access control is a vital component of cloud security. Typical controls include:

User Access Control: Implement system and application access controls that ensure only authorized users access cloud data and applications. A Cloud Access Security Broker (CASB) can be used to enforce access controls

Device Access Control: Block access when a personal, unauthorized device tries to access cloud data.

Malicious Behaviour Identification: Detect compromised accounts and insider threats with user behaviour analytics (UBA) so that malicious data exfiltration does not occur.

Malware Prevention: Prevent malware from entering cloud services using techniques such as file-scanning, application whitelisting, machine learning-based malware detection, and network traffic analysis.

Privileged Access: Identify all possible forms of access that privileged accounts may have to your data and applications, and put in place controls to mitigate exposure.

Compliance: Existing compliance requirements and practices should be augmented to include data and applications residing in the cloud.

Risk Assessment: Review and update risk assessments to include cloud services. Identify and address risk factors introduced by cloud environments and providers. Risk databases for cloud providers are available to expedite the assessment process.

Compliance Assessments: Review and update compliance assessments for PCI, HIPAA, Sarbanes-Oxley and other application regulatory requirements.

2.5 Importance of Cloud Security

According to recent research, 1 in 4 companies using public cloud services have experienced data theft by a malicious actor. An additional 1 in 5 has experienced an advanced attack against their public cloud infrastructure. In the same study, 83% of organizations indicated that they store sensitive information in the cloud. With 97% of organizations worldwide using cloud services today, it is essential that everyone evaluates their cloud security and develops a strategy to protect their data.¹ Cloud security from McAfee enables organizations to accelerate their business by giving them total visibility and control over their data in the cloud. Learn more about McAfee's cloud security technology solutions.

III. CONCLUSION

According to recent research, 1 in 4 companies using public cloud services have experienced data theft by a malicious actor. An additional 1 in 5 has experienced an advanced attack against their public cloud infrastructure. In the same study, 83% of organizations indicated that they store sensitive information in the cloud. With 97% of organizations worldwide using cloud services today, it is essential that everyone evaluates their cloud security and develops a strategy to protect their data. Cloud security from McAfee enables organizations to accelerate their business by giving them total visibility and control over their data in the cloud. Learn more about McAfee's cloud security technology solutions.

REFERENCES

- [1]. Boyang Wang (2014), Oruta: Privacy-Preserving Public Auditing for Shared Data in the Cloud, IEEE Transactions On Cloud Computing, Vol. 2, No. 1, January-March 2014, pp. 43-56.
- [2]. Shini.S.G(2012), Cloud Based Medical Image Exchange-Security Challenges, Procedia Engineering 38 (2012) pp. 3454 – 3461.
- [3]. KaipingXue(2018), Combining Data Owner-side and Cloud-side Access Control for Encrypted Cloud Storage, IEEE Transactions on Information Forensics and Security.
- [4]. CONG WANG(2013), Privacy-Assured Outsourcing of Image Reconstruction Service in Cloud, IEEE Transactions On Emerging Topics In Computing, Volume 1, No. 1, June 2013, pp. 166-177.
- [5]. Zhongbo Shi(2014), Photo Album Compression for Cloud Storage Using Local Features, IEEE Journal On Emerging And Selected Topics In Circuits And Systems, Vol. 4, No. 1, March 2014.
- [6]. RajkumarBuyya(2013), Introduction to the IEEE Transactions on Cloud Computing, IEEE Transactions On Cloud Computing, Vol. 1, No. 1, January-June 2013 2168-7161/13.
- [7]. Israna Hossain Arka(2014), Collaborative Compressed I-Cloud Medical Image Storage with Decompress Viewer, International Conference on Robot PRIDE 2013-2014 - Medical and Rehabilitation Robotics and Instrumentation, Conf. PRIDE 2013-2014, Procedia Computer Science 42 (2014) pp. 114 – 121.
- [8]. Sajida Karim(2020), The evaluation video quality in social clouds, Entertainment Computing 35 (2020) 100370.
- [9]. H. B. Kekre(2016), Color Image Compression using Vector Quantization and Hybrid Wavelet Transform, Twelfth International Multi-Conference on Information Processing-2016 (IMCIP-2016), Procedia Computer Science 89 (2016) pp. 778 – 784.
- [10]. Fouad KheliB(2018), Secure and Privacy-preserving Data Sharing in the Cloud based on Lossless Image Coding, Preprint submitted to Signal Processing February 13, 2018.
- [11]. Ranjeet Kumar(2019), An efficient technique for image compression and quality retrieval using matrix completion, Journal of King Saud University – Computer and Information Sciences.
- [12]. MamtaMeena(2016), Hybrid Wavelet Based CBIR System using Software as a Service (SaaS) Model on public Cloud, 7th International Conference on Communication, Computing and Virtualization 2016, Procedia Computer Science 79 (2016) pp. 278 – 286.
- [13]. B. Nivedha(2017), Lossless Image Compression In Cloud Computing, 2017 International Conference on Technical Advancements in Computers and Communications, 978-1-5090-4797-0/17.
- [14]. J. Smith(2012)15, Progressive encoding and compression of surfaces generated from point cloud data, Computers & Graphics 36 (2012) pp. 341–348.
- [15]. Man-Wen Tian (2019), Research on image recognition method of bank financing bill based on binary tree decision, J. Vis. Commun. Image R. 60 (2019) pp. 123–128.
- [16]. A.M. Vengadapurvaja (2017), An Efficient Homomorphic Medical Image Encryption Algorithm For Cloud Storage Security, 7th International Conference on Advances in Computing & Communications, ICACC-2017, 22- 24 August 2017, Cochin, India Procedia Computer Science 115 (2017) pp. 643–650.
- [17]. Chi Yang(2013), A spatiotemporal compression based approach for efficient big data processing on cloud, Journal of Computer and System Sciences.
- [18]. ChaoweiYang(2016), Utilizing Cloud Computing to address big geospatial data challenges, Computers, Environment and Urban Systems.

- [19]. Farhan IsrakYen(2019), Efficient Image Compression for Cloud System, 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), 24-25 December, 978-1-7281-6099-3/19.
- [20]. XingyueChen(2017), A Remote Data Integrity Checking Scheme for Big Data Storage, 2017 IEEE Second International Conference on Data Science in Cyberspace.